Critical appraisal of the medical literature

Jane Hirst
MBBS, MPH, PhD FRANZCOG
University of Oxford

Training course in research methodology and research protocol development
Geneva 2018
What makes a study reliable?
Nine out of ten women we asked believe that “Youth Code” makes their skin firmer and younger looking

Which design maximises the chance of getting the result the company wants?

a) Ask women buying “Youth Code” in the shops whether they agree their skin is firmer and younger looking?

a) Ask a random sample of women to try “Youth Code” and then comment on whether they agree that their skin is firmer and younger looking?
What is critical appraisal?

• Carefully and systematically evaluate research to assess:
  • Validity (is these findings trustworthy?)
  • Value (what do the results show?)
  • Relevance (How do these results relate to my clinical practice?)
Critical appraisal: a key component of evidence based medicine

- Clinical Problem
- Define the question
- Search for the evidence
- Critical Appraisal
- Decide what action to take
- Evaluate your new practice
Asking the right question

- **P**: Population
- **I**: Intervention
- **C**: Comparator
- **O**: Outcome
<table>
<thead>
<tr>
<th>Tips for Building</th>
<th>Patient or Problem</th>
<th>Intervention (a cause, prognostic factor, treatment, etc.)</th>
<th>Comparison Intervention (if necessary)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Starting with your patient, ask “How would I describe a group of patients similar to mine?” Balance precision with brevity.</td>
<td>Ask “Which main intervention am I considering?” Be specific.</td>
<td>Ask “What is the main alternative to compare with the intervention?” Again, be specific.</td>
<td>Ask “What can I hope to accomplish?” or “What could this exposure really affect?” Again, be specific.</td>
</tr>
<tr>
<td>Example</td>
<td>“In patients with heart failure from dilated cardiomyopathy who are in sinus rhythm …”</td>
<td>“… would adding anticoagulation with warfarin to standard heart failure therapy …”</td>
<td>“… when compared with standard therapy alone …”</td>
<td>“… lead to lower mortality or morbidity from thromboembolism. Is this enough to be worth the increased risk of bleeding?”</td>
</tr>
</tbody>
</table>
Choosing right study design

• Some study designs are not appropriate to answer certain questions

• All study designs are prone to different biases
Pyramid of evidence
So are RCTs the gold standard for evidence?

.....depends

Slides from: K Mahtani, CEBM Oxford
Limitations of RCTs

• Excellent vs Poor RCTs – quality varies
  • Impact on interpretation of result (external validity)?

• Expensive and time consuming
  • £250k - £millions over 2-5 years+

• May not always be the right study design to answer that question
A RCT to examine if smoking causes lung cancer

• 30 healthy Oxford Students
• Randomise to 2 groups
  • Gp1 smokes 20 cigarettes per day every day
  • Gp2 no smoking
Types of research

• What is the best study design for answering this type of question?
  • Aetiology
  • Diagnosis
  • Prognosis
  • Harm
  • Effectiveness
  • Qualitative
How to critically appraise an article

• **Validity**: methods to check that the biases for which that particular study design is prone have been minimised

• Results

• Clinical relevance
Validity

Internal

External
Bias

“the systematic deviation of the results of a study from the truth because of the way it has been conducted, analysed or reported”

Burls, “What is Critical Appraisal” 2009
## Sources of bias in clinical trials

<table>
<thead>
<tr>
<th>Bias Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>Biased allocation to comparison groups</td>
</tr>
<tr>
<td>Performance bias</td>
<td>Unequal provision of care apart from treatment under evaluation</td>
</tr>
<tr>
<td>Detection bias</td>
<td>Biased assessment of outcome</td>
</tr>
<tr>
<td>Attrition bias</td>
<td>Biased occurrence and handling of deviations from protocol and loss to follow up</td>
</tr>
</tbody>
</table>

Juni, BMJ 2001
Assessing Trials of effectiveness

Questions to ask:
1. Are the results of the trial valid?
2. What are the results?
3. Will the results help locally?

From: Critical Appraisal Skills Program, Oxford
https://casp-uk.net/
Checklists for clinical trials

CONSORT (Consolidated Standards of Reporting Trials) is a set of guidelines for the design, conduct, and reporting of research studies in the health sciences. It provides a standard way to report clinical trials to allow fair, valid, and comparable evaluations oftheir methodological quality and results.

The CONSORT 2010 checklist is designed to help authors report their clinical trial results in a way that is clear, complete, and transparent. It includes items that should be considered when reporting a clinical trial and provides a template for presenting the information.

SIGN (Scottish Intercollegiate Guidelines Network) is a guideline development organization that provides evidence-based clinical guidelines and best practice recommendations for healthcare professionals.

CEBM (Centre for Evidence-Based Medicine) is a research center located at the University of Oxford that conducts and promotes high-quality research to inform evidence-based medical decision-making.

CASP (Critical Appraisal Skills Programme) is a training program aimed at teaching healthcare professionals how to critically appraise clinical evidence. The goal is to help users make sense of the evidence by systematically evaluating its quality and relevance.
11 useful questions for critical appraisal of a randomised trial

1. Did the trial address a clearly focused issue?  
   □ Yes  □ Can’t tell  □ No

   Consider: An issue can be ‘focused’ in terms of
   • The population studied
   • The intervention given
   • The comparator given
   • The outcomes considered

Oxford Critical Appraisal Skills Programme (CASP)
Representative: Are the trial subjects representative of patients in this setting?

All people in study setting

Eligible participants

In trial
2. Was the assignment of patients to treatments randomised? □ Yes  □ Can’t tell  □ No

Consider:
- How was this carried out, some methods may produce broken allocation concealment
- Was the allocation concealed from researchers?
Why randomise?

• Minimises measured and unmeasured confounding
Minimising allocation bias

- Centralised computer randomisation the best
- Other methods such as sealed envelopes doubtful
- Non randomised: date of birth, alternate patients alternate days, etc.
If answer to first two questions is no....
Detailed questions

3. Were patients, health workers and study personnel blinded?

Consider:
- Health workers could be; clinicians, nurses etc
- Study personnel – especially outcome assessors
4. Were the groups similar at the start of the trial?

- Yes
- Can’t tell
- No

Consider: Look at

- Other factors that might affect the outcome such as age, sex, social class, these may be called baseline characteristics
Maintenance: Were the groups treated equally?

5. Aside from the experimental intervention, were the groups treated equally?
6. Were all of the patients who entered the trial properly accounted for at its conclusion?

Consider:
- Was the trial stopped early?
- Were patients analysed in the groups to which they were randomised?
(B) What are the results?

7. How large was the treatment effect?
Consider:
- What outcomes were measured?
- Is the primary outcome clearly specified?
- What results were found for each outcome?
- Is there evidence of selective reporting of outcomes?

8. How precise was the estimate of the treatment effect?
Consider:
- What are the confidence limits?
- Were they statistically significant?

From: Critical Appraisal Skills Program, Oxford
https://casp-uk.net/
Intention to treat

• Once a participant is randomised, they should be analysed to the group they were assigned to

• Pros
  • Reflects “real life” e.g. non compliance
  • Unbiased estimate of true effect
  • Maintains sample size thus maintaining statistical power

• Cons
  • Noncompliance provides little data on efficacy
  • Treatment effect may be conservative
  • Dropouts/non-compliant/compliant subjects are different

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3159210/
What does this study tell us?

- **P values** (hypothesis testing):
  - Tests to exclude the null hypothesis

- **Confidence intervals** (estimation of effect)
  - Range of values within which the true effect is likely to lie
  - Wider the confidence interval, less precision in result

- **Relative Risk**
- **Absolute Risk**
- **Odds Ratios**
- **Number needed to treat**
(C) Will the results help locally?

9. Can the results be applied in your context?
(or to the local population?)

Consider:

- Do you have reason to believe that your population of interest is different to that in the trial
- If so, in what way?

☐ Yes  ☐ Can’t tell  ☐ No
10. Were all clinically important outcomes considered?

Consider:

- Is there other information you would like to have seen?
- Was the need for this trial clearly described?

11. Are the benefits worth the harms and costs?

Consider:

- Even if this is not addressed by the trial, what do you think?
Conclusion

• Critical appraisal helps us decide whether evidence is valid, what the results tell us and whether the study is relevant to our setting

• Checklists are available to help

• Don’t believe everything you read in journals!